

Hydrophobicity Variations along the Surface of the Coiled-Coil Rod May Mediate Striated Muscle Myosin Assembly in *Caenorhabditis elegans*

Pamela E. Hoppe and Robert H. Waterston

Department of Genetics, Washington University School of Medicine, St. Louis, Missouri 63110

Abstract. *Caenorhabditis elegans* body wall muscle contains two isoforms of myosin heavy chain, MHC A and MHC B, that differ in their ability to initiate thick filament assembly. Whereas mutant animals that lack the major isoform, MHC B, have fewer thick filaments, mutant animals that lack the minor isoform, MHC A, contain no normal thick filaments. MHC A, but not MHC B, is present at the center of the bipolar thick filament where initiation of assembly is thought to occur (Miller, D.M., I. Ortiz, G.C. Berliner, and H.F. Epstein. 1983. *Cell*. 34:477–490). We mapped the sequences that confer A-specific function by constructing chimeric myosins and testing them in vivo. We have identified two distinct regions of the MHC A rod that are sufficient in

chimeric myosins for filament initiation function. Within these regions, MHC A displays a more hydrophobic rod surface, making it more similar to paramyosin, which forms the thick filament core. We propose that these regions play an important role in filament initiation, perhaps mediating close contacts between MHC A and paramyosin in an antiparallel arrangement at the filament center. Furthermore, our analysis revealed that all striated muscle myosins show a characteristic variation in surface hydrophobicity along the length of the rod that may play an important role in driving assembly and determining the stagger at which dimers associate.

MYOSIN is the motor that drives all actin-based motility. In recent years, an increasingly diverse array of myosins have been identified that participate in a variety of cellular processes (for review see Mooseker and Cheney, 1995). Myosin II is the conventional double-headed myosin. Cytoplasmic myosin IIs participate in dynamic processes, such as cytokinesis and morphogenesis, that require the assembly of transient myosin minifilaments in different subcellular locations at specific times during the cell cycle or during development. In contrast, striated muscle myosin contributes to a highly ordered filament lattice that is important for efficient contraction. We are interested in the mechanisms by which myosin is assembled into the thick filament structure. The identification and characterization of functional domains within the myosin heavy chain (MHC)¹ molecule is important for the eventual understanding of the development, action, and functional regulation of muscle. In addition, these studies may contribute to the understanding of the molecular events that are required for assembly of other molecules, including other types of myosin, into ordered structures.

The myosin heavy chain molecule is a large, multifunctional protein (see Fig. 1 *a*). The NH₂ terminal globular portion contains actin-binding and ATPase activities that move thin filaments past thick filaments to yield muscle contraction, and is the site of light chain association. The rod, where the two heavy chains dimerize by forming an α -helical coiled coil, is the structural component that is responsible for the assembly of myosin into thick filaments. In addition, some myosins, including *Caenorhabditis elegans* body wall muscle myosins, have a COOH-terminal nonhelical region containing phosphorylation sites (Castellani et al., 1988). The other major structural protein in the thick filaments of *C. elegans* and other invertebrates is paramyosin, a coiled-coil protein homologous to the COOH-terminal three fourths of the myosin rod (Fig. 1 *a*) (Kagawa et al., 1989). Paramyosin and associated minor proteins form the core of the thick filament (Deitiker and Epstein, 1993) upon which myosin assembles.

The two MHC isoforms that coassemble in *C. elegans* body wall muscle have specialized to play functionally distinct roles in filament assembly. Using isoform-specific antibodies, Miller et al. (1983) showed that the minor isoform, MHC A, forms the central 2 μ m of the 10- μ m-long bipolar thick filament. This central 2 μ m includes at its center the bipolar region of the thick filament where myosin molecules associate in an antiparallel (tail-to-tail) fashion, as well as extensive polar regions where myosin molecules associate in parallel fashion, and is the site at which

Please address all correspondence to Pamela E. Hoppe, Department of Genetics, Washington University School of Medicine, BOX 8232, 4566 Scott Avenue, St. Louis, MO 63110. Tel.: 314-362-2722; Fax: 314-362-2985.

1. Abbreviation used in this paper: MHC, myosin heavy chain.

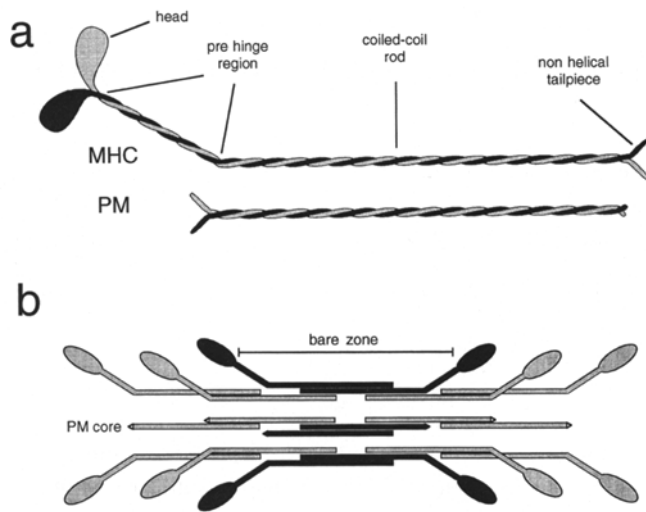


Figure 1. The major proteins of the invertebrate thick filament are myosin and paramyosin. (a) Schematic diagrams of the two major structural proteins of the thick filament, MHC and paramyosin (PM), with NH₂-termini to the left. Both proteins form dimers through the formation of a long coiled-coil domain or "rod." The complete myosin complex is formed by association of the light chains with the globular head, which contains the motor activities. The COOH-terminal, nonhelical tailpiece of *C. elegans* MHC contains three potential phosphorylation sites, based on the known *in vivo* phosphorylation of the similar NH₂-terminal region of paramyosin (Schriefer and Waterston, 1989). Apart from the short, nonhelical, NH₂-terminal segment, paramyosin is largely a coiled coil, and by sequence identity, aligns with the COOH-terminal three fourths of the myosin rod, as shown here; the corresponding portion of the MHC rod is in effect the light meromyosin sequences that form the backbone of the thick filament (for review of myosin structure see Squire, 1981). Based on the "hinge" definition of Lu and Wong (1985), the NH₂-terminal portion of the MHC rod that is not contained in the shorter paramyosin molecule, the "prehinge region," is part of heavy meromyosin subfragment 2 (S2). (b) Schematic diagram of a longitudinal section of the central region of the bipolar thick filament. The simplified molecules represent the dimeric paramyosin protein, and the MHC complex with associated light chains. The precise overlaps of the molecules assembled into the thick filament are unknown. Myosin and paramyosin must associate in an antiparallel or tail-to-tail fashion in the center of the filament (black), whereas elongation of the two filament arms occurs by parallel addition (gray). NH₂ termini are indicated by the myosin head for MHC and as a pointed end for paramyosin. In *C. elegans*, only MHC A is present in the central region. The "bare zone" refers to that central portion of the filament devoid of myosin heads.

filament initiation is thought to occur. The major isoform, MHC B, forms the remainder of the thick filament by adding in a parallel manner to either side of this central region (Fig. 1 b). The paramyosin-containing core extends almost the entire length of the filament.

Genetic analysis has confirmed the functional differences suggested by the differential subcellular localization of the two isoforms, and has demonstrated that MHC A is uniquely capable of thick filament initiation. Many mutations in the *unc-54* locus, the structural gene for MHC B (Epstein et al., 1974; Waterston et al., 1982), have been isolated in screens for uncoordinated animals. Animals ho-

mozygous for null mutations in *unc-54* are paralyzed, though viable and fertile, and exhibit disorganized body wall muscle containing a reduced number of thick filaments. These filaments can be of normal length, and they contain MHC A along their entire length (Epstein et al., 1986). In contrast, *myo-3* alleles (mutations that eliminate MHC A) cause embryonic paralysis and death due to the absence of a functional myofilament lattice, and mutant animals contain no normal thick filaments (Waterston, 1989). Therefore, despite that fact that it comprises only 20–30% of body wall myosin, MHC A is required for viability, and apparently plays a unique and essential role in the initiation of thick filament assembly. The unique function of MHC A has been confirmed by transformation experiments (Fire and Waterston, 1989) in which introduced copies of *myo-3* (MHC A) were able to improve the motility of *unc-54* mutants, consistent with results obtained earlier with chromosomal duplications (Riddle and Brenner, 1978; Waterston et al., 1982; Otsuka, 1986; Miller and Maruyama, 1986). In the reciprocal test, however, additional copies of the *unc-54* gene were unable to rescue *myo-3* lethality.

Paramyosin is encoded by the *unc-15* gene (Waterston et al., 1977). Mutant animals that lack paramyosin are paralyzed, and contain abnormally short and wide thick filaments (Mackenzie and Epstein, 1980) in which the distinct localization of MHC A and B is lost (Epstein et al., 1986).

The absence of the major isoform, MHC B, at the filament center, as well as the lethality associated with mutations that eliminate MHC A, suggest that the B isoform has specialized and apparently lost the ability to participate in some aspect of filament formation thought to be common to all myosins. The localization of the minor isoform, MHC A, to the center of the bipolar thick filament, where initiation of assembly may occur, as well as the genetic evidence that normal thick filaments are absent in mutants lacking MHC A, suggest that "A-specific function" is required in the early steps of filament formation. MHC A and MHC B protein sequences are 65% identical, and comparisons (Dibb et al., 1989) do not readily identify features or regions particularly different between the two that might be responsible for the unique function of MHC A.

The functional specialization of the two MHC isoforms that coassemble in *C. elegans* body wall muscle provides a unique opportunity to probe *in vivo* the molecular events involved in the initiation of thick filament assembly in striated muscle. To elucidate MHC A-specific function and thereby learn about the molecular events of thick filament initiation, we used the molecular genetics of *C. elegans* to map the sequences that are important for MHC A function *in vivo*. Our results identify regions of the rod that are important for MHC A function. Detailed comparisons of the differences between isoforms A and B in these two regions provide new insights into the characteristics of coiled coils that are important for assembly of myosin dimers into the highly ordered structure of the thick filament. In addition to electrostatic interactions analyzed in earlier work (Parry, 1981; McLachlan and Karn, 1982), our results suggest that characteristic variations in hydrophobicity on the rod surface also play a role in determining which dimer-dimer interactions occur *in vivo*.

Materials and Methods

Worm Cultures

General methods for culturing and handling *C. elegans* have been described (Brenner, 1974; Wood, 1988). N2 wild type was the parent of all strains used in this work. Strains were obtained from the Cambridge collection MRC Laboratory of Molecular Biology (Cambridge, England) and the Caenorhabditis Genetics Center (University of Minnesota St. Paul, MN).

Injection

Transformed lines carrying extrachromosomal arrays were generated as described (Mello et al., 1991) using a 50:1 ratio of *rol-6*/myosin DNA at 200 ng/ μ l in 10 mM Tris, 1 mM EDTA, pH 8. Each construct was tested for the rescue of *myo-3(st378)* by injecting DNA into heterozygous animals [*myo-3(st378)*/*eDf1*] that carry a chromosome bearing a *myo-3* mutation linked to a marker gene *sma-1* (small), balanced by the *eDf1* chromosome (which suppresses recombination in the *myo-3* interval). The F1 progeny were scored both for rescued small animals and for rolling heterozygotes. If small animals were obtained, rescue was confirmed by observing their offspring: unstable transformation events yield hermaphrodites that give rise to only mutant progeny, thus confirming that rescue, rather than recombination, occurred. The F1 rolling heterozygotes were picked to establish heritably transformed lines. If *myo-3* rescue was not obtained, the negative result was confirmed by establishing lines that rescue *unc-54*, crossing these arrays into a balanced *myo-3* background to score for rescue, and then crossing back into *unc-54* to confirm that the arrays retained activity.

To assay chimeras for general myosin function, we injected animals homozygous for either a null allele (*e190*) or a temperature-sensitive allele (*e1301*) of *unc-54*. For *e1301*, hermaphrodites grown at 15°C were injected. Eggs were collected at 15°C, and then moved to the restrictive temperature (25°C) to score for rescue.

Wild-type Genomic Clone Cassettes

A conserved unique *Sna*BI site was used to divide both genes into genomic cassettes which were made in a modified pAST vector (Fire, 1986), pAST25/35, with polylinker (destroyed *Eco*RI), *Sal*I, *Eco*RV, *Sna*BI, *Apa*I, (destroyed *Hind*III), from annealed 880 AGCTGGGCCCTACGTAGATATCGTCGAC/ A038 AATTGTGCGACGATATCTACGTAGGGCC. *unc-54* genomic DNA was recovered from the λ phage used for initial sequencing (McLachlan and Karn, 1982). In clone pJK24a, the 2.1-kb 5' *unc-54* *Sna*BI fragment was cloned into *Eco*RV-*Sna*BI-cut vector, regenerating the *Sna*BI site in the coding region. Similarly, the 7.0-kb 3' *Sna*BI-*Hpa*I *unc-54* fragment was cloned into the *Sna*BI-cut vector to make clone pJK28b. The genomic DNA used to determine the *myo-3* sequence (Dibb et al., 1989) was obtained as λ clones, and analogous cassettes were constructed: pJK23a contains the 3.8-kb 5' *Nae*I-*Sna*BI fragment, and pJK26a contains a 7.0-kb *Sna*BI-*Acl*I fragment. The complete gene is made by opening the 5' cassette with a *Sna*BI-*Apa*I digest and inserting the 7.0-kb 3' *Sna*BI-*Apa*I fragment. We moved the 5' cassettes into pUC-19 containing the pAST 25/35 polylinker to make clones pPH23aP1 (*myo-3*) and pPH24aP1 (*unc-54*).

Chimeric Myosin Constructs

All portions generated by PCR and all chimeric borders were verified by DNA sequencing unless indicated otherwise. Bases italicized in oligonucleotides do not match the wild-type sequence.

Chimeras 1 and 3. The *myo-3* *Xba*I site formed by the S795 and R796 codons served as junction site in both chimeras. The *Xba*I site was introduced into the *unc-54* clone pJK28b (A794 to S, T795 to R) using two sets of oligos: 92 GGAGCCATGGCCAAGGGTCTACTCC/102 CCT-ATCTAGAGAGCTTCTCGTCACGGAT (adds *Xba*I site); and 91 CGA-GTCTAGAATCCTCACCGGATTCCTCAA (adds *Xba*I site)/ 90 AAGTC-CACCGAGTTGGTCTCTCGTGGT. The PCR products generated were cut with *Xba*I and either *Nco*I or *Bst*XI, and were ligated together into *Nco*I-*Bst*XI-cut pJK28b. The resulting clone was confirmed by heteroduplex mismatch analysis (see below) and DNA sequencing through ligation sites. In the *myo-3* subclone, an *Xba*I site 830 bp 3' to the putative polyadenylation site was destroyed, deleting 80 bp: pJK26a was digested with *Xba*I and *Apa*I, and the two large fragments were ligated with annealed 97 CGGGTCCGGA/98 CTAGTCGGGACCCGGGCC. This yielded sub-

clones pPHU-9 (*unc-54*) and pPHM-72 (*myo-3*) containing unique *Xba*I sites at homologous positions. Complete genes were made by inserting *Sna*BI-*Apa*I fragments of pPHM-72 and pPHU-9 into pJK23a and pJK24a, respectively, to generate constructs pPH23M4 (*myo-3*) and pPHCU3 (*unc-54*). Clones pPHU-9 and pPHM-72 enabled a swap of restriction fragments to generate subclone pPHUM7, containing *unc-54* sequences from *Sna*BI to *Xba*I and *myo-3* *Xba*I to *Apa*I, and subclone pPH26aU9, containing *myo-3* from *Sna*BI to *Xba*I and *unc-54* from *Xba*I to *Apa*I. The pPH26aU9 *Sna*BI-*Apa*I fragment was inserted into pJK24a to complete chimera 1. The pPHUM7 *Sna*BI-*Apa*I fragment was inserted into pJK23a to complete chimera 3. Each construct was moved into Bluescript (SK-) using a *Sal*I-*Apa*I digest.

Chimera 2. The chimeric junction is a conserved *Cl*aI site (*unc-54* codons I1078 and D1079; *myo-3* codons I1079 and D1080). The construct was made in pUC-19 through several subcloning steps that resulted in the union of *myo-3* fragments pPH23aP1 *Sal*I-*Sna*BI, pJK26a *Sna*BI-*Xba*I, and pJK26a *Xba*I-*Cl*aI with an *unc-54* *Cl*aI-*Apa*I fragment from chimera 1.

Chimeras 4 and 6. The 2.2-kb *Sph*I-*Apa*I fragment from pPHUM7 was replaced by SOE PCR. PCR 1: B0153 CGCATCCAAGAAAAGGAA/928 TGATCGCATCTTTGACAGGGCATTTTC on pJK26a. PCR 2: 929 CTGTCAAAGATGCGATCAAAATCCCGT/679 GGGCAGTGAGCGCAACGC (pUC) on pJK28b. The *Sna*BI-*Apa*I fragment of the resulting clone was inserted into pPH24aP1. A clone with a premature stop (*unc-54* codon K1935 to TAA) was inserted into pPH24aP1 to make chimera 6 (the original stop was changed to TAG).

Chimera 5. The 1.6-kb *Mlu*I-*Apa*I fragment of pJK28b was replaced by SOE PCR. PCR 1: A037 GCTCGAACACGCGTGCG/A039 CT-TGTTACGCATCTTTGAGAGGGAGTT on pJK28b. PCR 2: A034 TCA-AAGATGCGTAAACAAGATTCGTGCA/679 on pJK26a. The *Sna*BI-*Apa*I fragment was inserted into pPH24aP1. DNA sequencing revealed a change in L1929 (to CTT).

Chimera 7. The 1.1-kb *Sph*I-*Eag*I fragment from chimera 6 was replaced with a pJK28b PCR product: B0601 CTTGAAAAGCATGCAG-GCTTCCCTCGAGACC (adds an *Sph*I site and destroys a *Hind*III site, no coding changes)/A731 GGTGGGAGCACAGGGAG.

Chimera 8. The 854-bp *Xba*I-*Cl*aI pPHCU3 fragment was ligated into *Xba*I-*Cl*aI-cut chimera 7. The *Xba*I-*Apa*I fragment from the resulting clone was inserted into *Xba*I-*Apa*I-cut pPHUM7. The *Sna*BI-*Apa*I fragment from the resulting clone was inserted into pPH24aP1.

Chimera 9. A polylinker of annealed A411 AGCTGGGCCGTC-GACGCATGC/A412 AATTGCATGCGTCGACGGGCC was ligated into *Hind*III-*Eco*RI-cut Bluescript SK-, and the 2.4-kb *Xba*I-*Sph*I chimera 11 fragment was cloned and transformed into CJ236 cells. Using the method of Kunkel et al. (1987), a *Bst*XI site was added using D0964 CGCCGCCAGGCCAAGAATTTGGAACACGAAGCCG (*unc-54* sequence AKNYQ to *myo-3* sequence AKNLE); D0965 CGAGAAC-CAACTCGGAGGACTTCG was used to destroy a *Bst*XI site (G1202 codon to GGA). The 7.0-kb *Sna*BI-*Apa*I chimera 7 subclone was cut with *Xba*I and *Sph*I to accept the altered subclone, yielding clone pPHsubXBS. From this, the 4.3-kb *Sna*BI-*Bst*XI fragment was removed and inserted into *Sna*BI-*Bst*XI-cut chimera 7.

Chimera 10. The 3.2-kb *Bst*XI-*Apa*I chimera 8 fragment was replaced with *unc-54* sequences from clone pPHsubXBS (see chimera 9). This construct failed to rescue *unc-54* animals. The complete *unc-54* construct pPHXBS (made by inserting pPHsubXBS *Sna*BI-*Apa*I into pPH24aP1) also failed to rescue, indicating that a mutation had occurred. We mapped the mutation by coinjecting overlapping restriction fragments, with vector sequences intact, from pPHXBS and wild-type *unc-54*, relying on in vivo recombination to form a complete gene during extrachromosomal array formation. The fragments were injected into *e190* homozygotes at a total DNA concentration of 150 ng/ μ l, with the 5' fragment as 0.5%, the 3' fragment as 10%, and *rol-6* marker as 90% of total DNA. The first set of injections combined a 5' *Apa*I-*Afl*III fragment with a 3' *Bam*HI-*Sal*I fragment. Rescue was obtained with pPHXBS *Apa*I-*Afl*III fragment/wild-type *Bam*HI-*Sal*I fragment, but not wild-type *Apa*I-*Afl*III fragment/pPHXBS *Bam*HI-*Sal*I, indicating the mutation lay 3' to the *Afl*III site at 5983. The next set combined a 5' *Apa*I-*Sph*I fragment with a 3' *Asp*I-*Sal*I fragment. Both combinations of fragments produced rescuing arrays, mapping the mutation to the region of overlap between the two fragments, *unc-54* 6410-6949. DNA sequencing revealed a single basepair deletion (G6751) in the R1544 codon. To generate a functional construct, we coinjected restriction fragments (the *Apa*I-*Sph*I fragment of nonfunctional chimera 10, containing the entire chimeric portion, and a wild-type *Asp*I-*Sal*I fragment) into *e190* animals and relied on in vivo recombination. Seven rescued lines were obtained; two of these arrays were crossed into a *myo-3* background to test for rescue. The correct join was detected by single ani-

mal PCR (Williams et al., 1992) across the recombination border using F0219 ACTCTCTCACCGCTTG(*myo-3*)/F0648 CCTTTCCTTAGCTTCG (*unc-54*).

Chimera 11. The 2.4-kb XbaI-SphI pPHUM7 fragment was replaced by the pJK28b PCR product 91/ B0417 GAAGGGCATGCTTTCAA-GAGCGGGCGGTG (adds an SphI site, no coding change). The SnaBI-ApaI fragment from this was ligated into pPH24aP1.

Chimera 12. The 1.1-kb SphI-EagI chimera 11 fragment was replaced by SOE PCR. PCR 1: A298 CCAGGTTGAAGTTTCCC/F0562 GCGC-TCTCAGCGTTCTTGAGCTCGTTGGC on chimera 11. PCR 2: F0561 GAGCTCAAGAACGCTGAGGAGCGCTCCAAG/A731 on pJK28b.

Chimera 13. An SphI-ApaI subclone of the chimera 11 COOH terminus was cut with SphI and HindIII to remove 0.8 kb, which was replaced by SOE PCR. PCR 1: B0599 TTGCTTGTTCACGGCGCCTGTG-TACTCGTTGAG/ B0601 on pJK28b. PCR 2: B0600 CTCAACGAGTCAAGGCCCGTTGAACAAGGACAA/ 928 on pJK26a. From this, the SphI and ApaI fragment was introduced back into chimera 11.

Chimera 14. The SphI-ApaI subclone of the chimera 13 COOH terminus was cut with BssHII and HindIII to remove 0.7-kb, which was replaced by SOE PCR. PCR 1: A299 GAACTGAAGAGATACC/A732 CTCTCGGTATCCTGGAAGCGCGCTTG on pJK28b. PCR 2: B0702 TTCCAGGATACCGAGAAGAAGTGGAG/928 on pJK26a. The resulting subclone was cut with SphI and ApaI, and introduced back into chimera 13.

Chimera 15. The 2.3-kb SphI-ApaI subclone of chimera 13 was cut with BclI and EagI to remove 0.8 kb, which was replaced by SOE PCR. PCR 1: A299/B0704 CTTGTTGGCGTCTTGATGACGTCTCTG on chimera 13. PCR 2: B0703 CATCAAGACGCCAACAAGAACCTCGG/A731 on pJK28b. This was replaced in chimera 13.

SOE PCR. We used the method of Horton et al. (1989) to create chimeric junctions. In PCRs 1 and 2, fragments with complementary ends are amplified in separate reactions. In PCR 3, these gel-purified products are mixed together with the two flanking oligonucleotides to form the join.

Heteroduplex Mismatch Analysis

Oligo pairs 324 CCCAAATGCCATCTGATTTG/325 GCGAACAA-GTCAATTGGG and 330 GCTCATCTTGAAGACATCCG/331 CTC-AGCGACAGCGTTCGGTGT were used to generate PCR fragments covering the NcoI-BstXI region. Wild-type or U-9 PCR fragments were end labeled (Maniatis et al., 1982), mixed with the corresponding unlabeled fragment at a 1:10 ratio, boiled for 5 min, and left under mineral oil at 65°C overnight in 0.3 M NaCl, 0.1 M Tris-Cl, pH 8, and 0.1 mM EDTA. Mismatches were detected as described in Cotton et al. (1988) and Montandon et al. (1989).

Sequencing

Double-stranded templates were sequenced using reagents and methods in the Sequenase Kit (U.S. Biochemical Corp., Cleveland, OH) or Taq Dye Deoxyterminators Cycle Sequencing Kit (Applied Biosystems, Inc., Foster City, CA).

Nucleotide and Protein Sequences

GenBank/EMBL/DDBJ accession numbers for *C. elegans* protein files are as follows: MHC A, P12844; MHC B, P02566; paramyosin, P10567 (however, the NH₂ terminus is incorrect, see Kagawa et al. [1989] for correct sequence). GenBank/EMBL/DDBJ accession numbers for DNA files are as follows: *unc-54*, J01050; *myo-3*, X08067. The 1,088-residue rod sequences from the following striated muscle myosins were aligned without gaps beginning at the rod-initiating proline: *Drosophila*, M61229; scallop, P24733; chicken fast skeletal, PIR JX0178; human perinatal, M36769; rat skeletal, A24922.

Statistical Analysis

To perform the equivalent of a paired, two-tailed *t* test, we calculated the hydrophobicity difference (*MHC A residue hydrophobicity* - *MHC B residue hydrophobicity*) for each coat residue in the rod. Positive values represent residues where A is more hydrophobic than B. The average difference was computed for each rod region. 95% confidence intervals were generated using a bootstrapping technique (Efron and Gong, 1983). The difference scores were resampled with replacement to generate 1,000 ranked values; the highest and lowest 25 values were deleted to define the 95% confidence interval. Confidence intervals for the average in regions

1 and 2, but not other rod regions, are positive and do not contain 0; thus, only regions 1 and 2 are significantly more hydrophobic in MHC A than in MHC B.

Results

MHC A-specific Function Maps to the Rod

We have used chimeric constructs in an *in vivo* assay to identify regions of the MHC that are important for A-specific function. The constructs have used genomic DNA fragments, thus testing native regulatory sequences as well as protein-coding portions of the gene. The *in vivo* assays involved introducing the recombinant genes into *C. elegans* as extrachromosomal arrays along with a marker construct to allow the identification of transformed animals independently of myosin construct function. Activity was determined in an *unc-54* loss-of-function background to assure that a functional MHC was produced, and in a *myo-3* null background to assess the MHC A-specific activity (see Materials and Methods for details).

The first set of chimeric constructs, illustrated in Fig. 2, involved substituting large portions of A sequence in the B gene or vice versa, and allowed us to map A-specific function to the α -helical coiled-coil rod. Chimera 1, which contains sequences coding for most of the A head with the remainder of the protein specified by B sequences, and chimera 2, which contains the entire A head and most of the preheing region of the rod (see Fig. 1 a), failed to rescue *myo-3* mutations, but did rescue *unc-54*. The gene expression in these constructs must be driven by *myo-3* regulatory sequences since all known *unc-54* regulatory sequences have been replaced in these constructs. A third construct, chimera 3, containing the B head and the A rod and nonhelical tailpiece, rescued *unc-54* and *myo-3* mutations. Taken together, these results show that the sequences encoding the NH₂-terminal half of MHC A, including regulatory sequences as well as coding sequences for the globular head and the light chain-binding sites, are neither necessary nor sufficient for A-specific filament initiation function.

We then tested the role of the nonhelical tailpiece region, which we define here as M1934 (beginning of McLachlan and Karn [1982] zone 40) to the end of the coding region so as to include all three potential phosphorylation sites (Schriefer and Waterston, 1989), and therefore including a small part of what is formally the COOH terminus of the rod. Chimera 4, specifying B head, A rod, and B tailpiece rescues *myo-3*; by contrast, chimera 5 specifies B head, B rod, and only the tailpiece from A, and fails to rescue *myo-3* (it does rescue *unc-54*). Chimera 6, a construct that specifies the B head and A rod but has a stop codon introduced before the sequences specifying the tailpiece, rescues *myo-3* lethality. Transformants carrying this construct, however, show disorganized muscle structure, a property we are currently characterizing further. Taken together, these results show that the rod alone is both necessary and sufficient for A-specific activity, and that the tailpiece is not essential for thick filament initiation, but is required for normal filament structure.

To localize further the sequences important for A-specific function, we made chimeras in which only a portion of








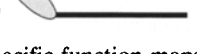
MHC A RESIDUES		RESCUE	
		<i>myo-3</i> FILAMENT INITIATION FUNCTION	<i>unc-54</i> GENERAL MYOSIN FUNCTION
		+	+
		-	+
CONSTRUCT			
Chimera 1	M1-R796 	-	+
Chimera 2	M1-I1079 	-	+
Chimera 3	S795-L1969 	+	+
Chimera 4	S795-M1933 	+	+
Chimera 5	M1933-L1969 	-	+
Chimera 6	S795-R1934* 	+	ND

Figure 2. MHC A-specific function maps to the rod. The MHC A sequences required for filament initiation function were mapped by constructing chimeric myosins of MHC A and MHC B, and testing these in vivo via germline transformation for their ability to substitute for the native MHC A and MHC B. For each chimera (1–6, *left column*) the precise MHC A amino acids included in each construct are indicated (*second column*) along with a schematic myosin molecule showing the portions of MHC A (*black*) and MHC B (*gray*). The last two columns summarize the results of in vivo rescue experiments testing the construct for filament initiation function (defined as ability to rescue *myo-3* mutations) and general myosin function (defined as ability to rescue *unc-54* mutations). The top two lines illustrate the results obtained by Fire and Waterston (1989) demonstrating that MHC A alone is able to rescue the lethality of *myo-3* mutations, while transgenic copies of either isoform can compensate for loss of *unc-54* gene activity. The six chimeric constructs demonstrate that the MHC A rod alone is necessary and sufficient for A-specific function. *R1934 is the last MHC A residue, but an additional N residue from MHC B is present at the COOH-terminus; ND, not done.

the MHC B rod was replaced by MHC A sequences. Our results, illustrated in Fig. 3, revealed that two distinct regions of the rod are independently sufficient to rescue *myo-3* mutations. The five constructs in the top portion of Fig. 3 define a 264-residue region in the central portion of the rod (region 1) that is sufficient to confer A-specific function. The five constructs drawn in the bottom portion of the figure define a 170-residue region at the COOH terminus of the rod (region 2) that is sufficient. Attempts to further refine the region failed: the two constructs chimera 14 and chimera 15, each of which contain one half of the 170-residue COOH-terminal region, did not contain full rescuing activity. Few of the lines obtained upon injection of *myo-3* heterozygotes produced rescued animals, and the few rescued animals were sickly and sterile. Because functional thick filaments are required for viability, the poor rescue may indicate that the chimeric myosin is unable to


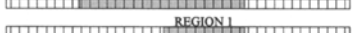



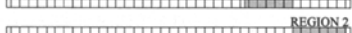


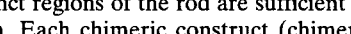
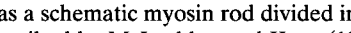
CONSTRUCT	ROD POSITIONS OF MHC A RESIDUES	RESCUE FILAMENT INITIATION
Chimera 7	1-761* 	+
Chimera 8	230-761 	+
Chimera 9	498-761 	+
Chimera 10	230-497 	-
Chimera 2	1-229* 	-
Chimera 11	762-1083 	+
Chimera 12	762-913 	-
Chimera 13	914-1083 	+
Chimera 14	999-1083 	P
Chimera 15	914-998 	P

Figure 3. Two distinct regions of the rod are sufficient for MHC A-specific function. Each chimeric construct (chimeras 2 and 7–15) is diagrammed as a schematic myosin rod divided into the 40 28-residue zones described by McLachlan and Karn (1982), with the NH₂ terminus to the left. MHC A sequences are shaded, and in the second column are assigned rod positions according to McLachlan (1984), where residue 1 is the rod-initiating proline, P851. The rod positions of the two smallest regions sufficient for A-specific function are boxed. The column at the far right indicates whether the construct was able to rescue *myo-3* mutations. P, partial function. All constructs dramatically improved the movement of *unc-54* mutants, and therefore specified functional myosins. (We did not attempt to quantify *unc-54* rescue further.) The top five constructs define region 1, a 264-residue region in the center of the MHC A rod (N1348-Q1611) that confers A-specific function in chimera 9. The bottom five constructs define region 2, a 170-residue region at the COOH terminus of the MHC A rod (A1764-M1933) that is independently sufficient for filament initiation (chimera 13). *These constructs contain MHC A residues NH₂-terminal to the rod that are not shown on the schematic diagram: chimera 7, S795-Q1611; chimera 2, M1-I1079.

promote formation of enough filaments, or that the filaments formed are not sufficiently stable or are otherwise only partially functional. The same result was obtained with lines that were selected through injection into *unc-54* and then crossed into a *myo-3* background to test for rescue. The partial activity of each construct, which was not seen in other constructs, suggests we have divided the active region. It is unclear without further experiments how widely the activity is distributed along the rod. Because these results indicate that the activity required for A-specific function must be a characteristic of a rod section rather than a single difference between isoforms A and B, we undertook a comparative sequence analysis of the two rod regions we had defined in the above chimeras to look for consistent differences in rod amino acid composition that correlate with the observed difference in function.

Comparison of Residues in the A and B Rods

In our analysis of the two sequences, we took advantage of the known structure of the myosin rod, a coiled-coil domain formed by the α helices of the two heavy chain monomers. Protein sequences that adopt a coiled-coil structure display a characteristic seven-residue repeat pattern, designated *a-g* (Fig. 4), where residues *a* and *d* form the core of the coiled coil, an internal hydrophobic seam. Other residues are next-to-core (*e* and *g*), or coat (*b*, *c*, and *f*).

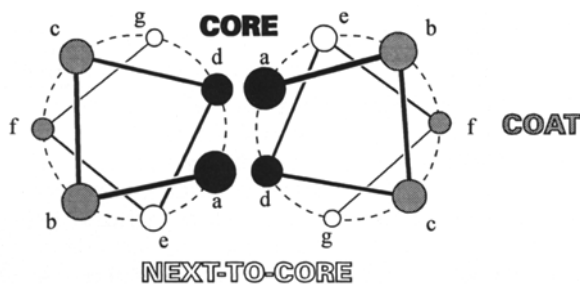


Figure 4. Schematic representing a cross-section of a coiled coil viewed from the NH₂ terminus. The positions of the residues within the heptad repeat, *a-g*, are shown for both helices. Positions *a* and *d* form the core, an internal hydrophobic seam bonding the two α helices together. Positions *e* and *g* fall next to the core, and often contain charged residues arranged so as to form salt bridges between the helices of the two monomers (McLachlan and Stewart, 1975). These ionic interactions are thought to strengthen the coiled coil, as well as dictate a parallel association of the two monomers. Coat positions *b*, *c*, and *f* are highly charged, and are exposed on the exterior of the rod, where they may mediate dimer-dimer interactions during assembly of the higher order structure of the thick filament. For reviews of coiled-coil literature see Cohen and Parry (1990, 1994).

Comparison of the two isoforms is most effectively done by considering residues at similar positions within the repeat.

Comparison of Charged Residues in MHC A and MHC B

Charge interactions are postulated to play a fundamental role in myosin assembly, both providing the driving force and determining the stagger at which two rods associate. Analysis of the distribution of basic and acidic residues reveals a 28-residue repeat unit or "zone" in which the strongest peak of positive charge is spaced 14 residues away from the strongest peak of negative charge. During the assembly of dimers into the thick filament, neighboring rods will associate at staggers that pair the peaks of positive and negative charge (Parry, 1981; McLachlan and Karn, 1982).

Because of the known importance of electrostatic interactions in the assembly of myosin dimers into the thick filament structure, we compared the content of charged residues in MHC A and B. The two isoforms are quite similar in total number of charges found in the rod (acidic: A = 249, B = 251; basic: A = 196, B = 200). To determine whether differential charge distribution could account for the differences in isoform function, we compared the location of charged residues in the rods of isoforms A and B. At each position within the rod (defined here as in Kagawa et al. [1989] as P851 to R1938, corresponding to zones 1 through part of zone 40 of McLachlan and Karn [1982]), we compared the charged residues in the two isoforms, and determined whether the charge at any one position was of the same or opposite sign. Of 414 positions in the 1,088-residue rod where both MHCs A and B contained a charged residue, charges of the same sign were found in 412 positions, while charges of opposite sign were found in only 2. "Unmatched" charges, those found at a particular position in only one isoform, totaled 28 in MHC A and 34 in MHC B. Many of these positions (Fig. 5), as well as the two positions that contained opposite charges, are

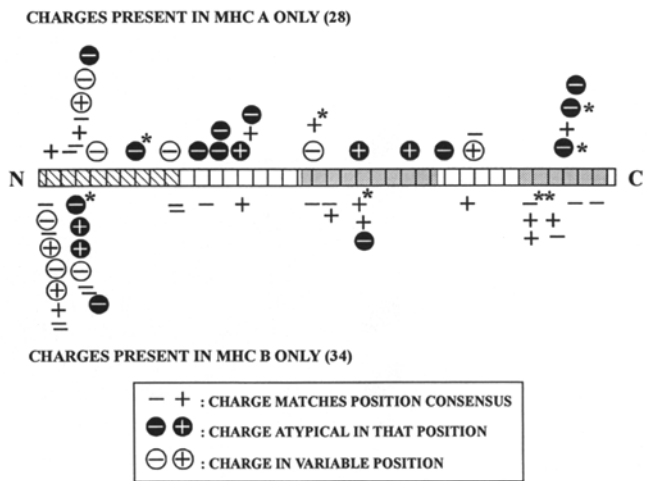


Figure 5. Distribution of the rod positions that contain a charged residue in only one of the two isoforms (unmatched charges), and comparison of these charges to the consensus derived from other striated muscle myosins. The myosin rod diagram is divided into the 40 28-residue zones described by McLachlan and Karn (1982), with the NH₂ terminus to the left. The two regions sufficient for A-specific function are shaded, and the prehing region (see Fig. 1 *a*) is hatched. Positions containing a charged residue in MHC A only are represented above the rod, those containing a charge in MHC B only are indicated below. Plain symbols, minus (-) or plus (+), indicate an acidic (D or E) or basic (K or R) residue that matches the consensus of striated muscle myosins (present in at least four of the five myosins chosen, see Materials and Methods). The symbols in the filled circles indicate atypical charges, those present in a position that is not charged in any of the other five myosins. The symbols in the open circles indicate charges in positions that are sufficiently varied in content that no consensus could be drawn. From this analysis, we found that the charge pattern in MHC A is more divergent: a higher proportion of unmatched charges in B are conserved in other myosins (25/34 for B vs. 8/28 for A), and MHC A has more unmatched charges in atypical positions (5/34 for B vs. 13/28 for A). Unmatched charges in next-to-core positions result in a different distribution of potential interhelical 1g-2e salt bridges (McLachlan and Stewart, 1975) within the coiled coil. A single asterisk indicates those charges that may participate in a favorable salt bridge, and a double asterisk indicates a charge that may participate in an unfavorable 1g-2e interaction. Region 2, and no other portion of the rod tested, shows a net difference in salt bridge number, where MHC A appears to have an increased number of favorable interhelical salt bridges that may result in a locally more stable helix with a lower charge density apparent at the rod surface.

found in the prehing region, the portion of the rod that is most variable (Dibb et al., 1989). The remaining charge differences are distributed fairly evenly along the rest of the rod, without an apparent concentration in the regions containing MHC A-specific activity (Fig. 5).

To determine how the differences in charge distribution in MHC A and B compare to other myosins, we examined the charge content in other myosins at the rod positions in which either A or B contained an unmatched charge. Most myosins are thought to be capable of both parallel and antiparallel assembly, and that is clearly the case for *Drosophila melanogaster*, where there is only one gene (alternative splicing does not affect the two regions of the rod defined in our study [George et al., 1989]). Our analysis re-

vealed that within the whole rod, and in both regions sufficient for A-specific function, the MHC A charge distribution is more divergent (Fig. 5). One model for A-specific function might have been that MHC B has specialized to perform only parallel assembly, and has thus lost a part of the charge pattern that is required only for antiparallel assembly. Because charge placement in MHC B is more similar to the consensus pattern, this model is an unlikely explanation for A-specific function.

The one aspect of charge distribution in MHC A that could be correlated with A-specific function was a reduction in surface charge. In both regions sufficient for A-specific function, the majority of unmatched charges in MHC B (5/6 and 5/7) are conserved residues found in coat positions, whereas few of the unmatched charges in the same regions of MHC A (1/4 and 1/4) fall into coat positions, resulting in a net loss of charge from the MHC A coat. Reduction of surface charge may also occur by a second mechanism in region 2, where unmatched charges in next-to-core positions result in a net gain in the number of favorable 1g-2e charge interactions within the MHC A coiled coil. Within region 2, MHC A loses one unfavorable interaction because of loss of a consensus charge, and gains two favorable salt bridges because of charges in atypical positions (Fig. 5). Because the additional intradimer salt bridges in MHC A are formed by atypical charges pairing with and therefore attenuating a conserved residue of opposite charge, we believe these changes may result in a reduction of the effective charge seen at the rod surface.

Differences in Hydrophobicity in Coat Positions

The overall content of hydrophobic residues (defined as those having a positive value on a hydrophobicity scale) is quite similar in the rods of MHC A and MHC B (Kyte and Doolittle, 1982: MHC A = 349, MHC B = 340; Engelman et al., 1986: MHC A = 470, MHC B = 466). Throughout the rod, however, the overall distribution of hydrophobic residues is not as well conserved between the two isoforms as the distribution of charged residues. Comparisons using the Kyte and Doolittle (1992) scale as a reference find 267 positions at which both isoforms have a hydrophobic residue, 82 positions that contain a hydrophobic residue in MHC A only, and 73 that contain a hydrophobic residue in MHC B only. Comparisons using the Engelman et al. (1986) scale find 394 positions in which both isoforms contain a hydrophobic residue, 76 with a hydrophobic residue only in MHC A, and 72 with a hydrophobic residue only in MHC B.

To investigate whether these differences were important for MHC A-specific function, we compared the hydrophobic content of the two isoforms in the three types of positions within the heptad repeat (core, next-to-core, and coat, see Fig. 4). The most striking difference in sequence composition was a statistically significant increase in the hydrophobicity of the MHC A coat residues only in those regions containing A-specific function (Fig. 6 b). Within both of these regions, the MHC A values are similar to those of paramyosin, for which a more hydrophobic coat has been correlated with its more internal position within the thick filament (Cohen et al., 1987).

To determine how the coat hydrophobicity values of

MHC A and B compare to those of other myosins, we performed the same calculations on selected striated muscle myosins from vertebrates and other invertebrates (Fig. 6 c). In both regions sufficient for function, but not in other rod regions, the MHC A value is the highest and the MHC B value is the lowest. This strengthens our conclusion that isoforms A and B differ significantly in coat hydrophobicity in these regions, and strongly suggests that this observed difference is important for A-specific function.

Our results suggest that differences in coat hydrophobicity in certain rod segments, defined by the borders chosen for our chimeric constructs, are important in at least one aspect of myosin assembly. To examine the variation in coat hydrophobicity in a less restricted fashion, we calculated hydrophobicity in a smaller window of sequence moved along the length of the rod (Fig. 7). This analysis revealed a characteristic hydrophobicity profile shared by MHCs A and B (Fig. 7 a). The regions where the two isoforms show the greatest difference in coat hydrophobicity are contained within the sequences we identified as important for MHC A-specific function. We then performed this same analysis on other myosins in the public databases, representatives of which are shown in Fig. 7 c, and found that the characteristic coat hydrophobicity profile is common to striated muscle myosins in vertebrates and invertebrates alike, suggesting that variations in hydrophobicity of the rod surface may play an important role in assembly of all myosins. Interestingly, when the hydrophobicity profiles for MHC A and paramyosin are compared, they are very similar overall, and particularly within the regions sufficient for MHC A-specific function (Fig. 7 b).

Because of the similarity of MHC A and paramyosin coat hydrophobicity scores, we examined other aspects of sequence composition that are known to differ between most myosins and paramyosin (Cohen et al., 1987; Kagawa et al., 1989). Paramyosin is thought to form a more rigid coiled-coil rod partly because of a reduced number of glycine residues. In addition, paramyosin sequences have a higher R/K ratio. When we compared isoforms A, B, and paramyosin in the region that all are homologous, we found that MHC A, like paramyosin, had fewer glycines (PM = 9, A = 15, B = 29) and a higher R/K ratio (PM = 1.1, A = 0.92, B = 0.83) than MHC B, indicating that the A rod is likely to be more rigid. Thus, in addition to its higher coat hydrophobicity, MHC A sequences have diverged along the length of the rod in ways that make it more similar to paramyosin.

Discussion

In our experiments, we have exploited a difference in the genetic properties of two MHC isoforms to elucidate aspects of the molecular basis for thick filament initiation. Using chimeric myosin constructs, we have mapped the sequences within MHC A that confer A-specific function *in vivo*. Neither the head nor the nonhelical tailpiece provide this activity; in fact, constructs that contain the MHC A rod and lack a tailpiece altogether rescue loss-of-function mutations in MHC A. Instead, the rescuing activity maps to two distinct regions of the coiled-coil rod, either of which is sufficient in chimeric myosins for the initiation function. Comparison of the MHC A and MHC B sequences

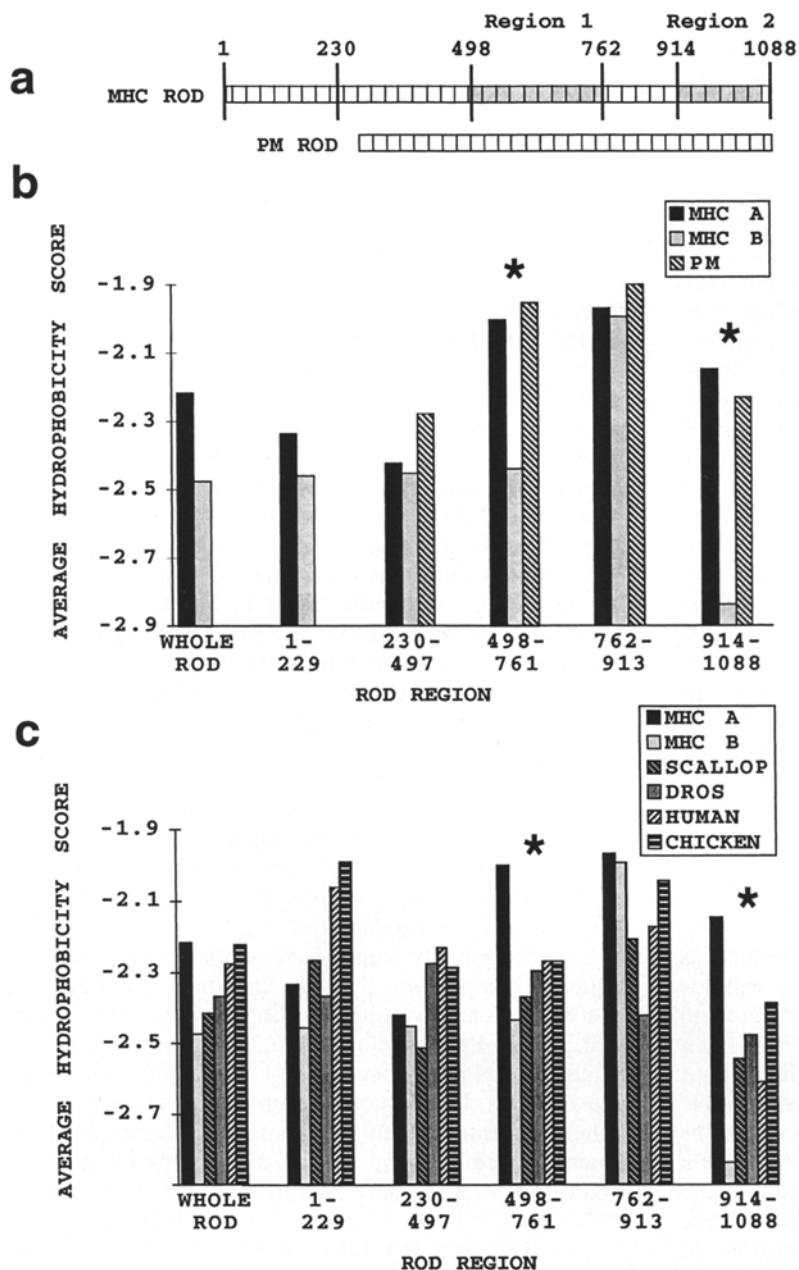


Figure 6. Comparison of the average coat hydrophobicity score in each region of the MHC rod tested independently for function *in vivo*. (a) Diagrams of the myosin and paramyosin rods divided into 28-residue zones and aligned to each other by homology, NH₂ termini to the left. The boundaries of the rod regions for which we computed the hydrophobicity scores are indicated by heavy vertical lines and assigned a rod position as in McLachlan (1984), where residue 1 is the rod-initiating proline. Regions of the rod that we have identified as sufficient for filament initiation function are shaded and labeled as region 1 or 2. (b) To calculate the average coat hydrophobicity score, each coat residue was weighted using the Kyte and Doolittle (1982) scale, and the sum of the scores was divided by the total number of coat positions included. Similar results were obtained using the Engelman et al. (1986) scale. Comparison of the average coat hydrophobicity in each rod region reveals that the hydrophobicity scores in both regions 1 and 2 (marked with asterisks), and not elsewhere in the rod, are significantly different in MHC A and MHC B ($P < 0.05$ in a paired *t* test using either the Kyte and Doolittle (1982) or Engelman et al. (1986) scales, see Materials and Methods). Within both regions, the MHC A scores are quite close to the scores of the homologous region in paramyosin. (c) Comparison of the hydrophobicity scores of MHC A and B to those of other striated muscle myosins from vertebrates and invertebrates. In both regions 1 and 2, and nowhere else in the rod, the MHC A value is the highest and the MHC B value is the lowest, confirming that the two isoforms differ significantly in hydrophobicity, and that the increased values in MHC A are found only in regions that are important for A-specific function.

within these two regions revealed a marked difference in the hydrophobicity of the coat residues, those exposed on the rod surface where they can mediate protein-protein interactions during assembly. We propose that sequences within these regions play an important role in the intermolecular interactions that are critical for the initiation of thick filament assembly. Furthermore, analysis of the coat positions along the length of the entire rod revealed a characteristic hydrophobicity profile that is conserved in all striated muscle myosins. We propose that this variation in coat hydrophobicity plays an important role in thick filament assembly.

Differences in Coat Hydrophobicity Correlate with Differences in Assembly Behavior

Within the two regions sufficient for A-specific function,

the coat hydrophobicity of MHC A is higher than that of MHC B and closely approximates that of paramyosin, a molecule that can be viewed as a fragment of myosin that has specialized to perform a strictly structural role in the formation of the large thick filaments that are present in invertebrate muscle. In addition, other aspects of the MHC A rod sequence, such as glycine content, are more similar to paramyosin. These sequence similarities, particularly within the critical regions, suggest that MHC A has specialized structural roles that may be shared with paramyosin. MHC B may be specialized for other structural roles, such as elongation in parallel assembly.

Our observation that the MHC A rod sequence shares characteristics in common with the paramyosin sequence correlates well with the similar *in vitro* behavior of MHC A and paramyosin. Biochemical experiments in which native thick filaments were dissociated by treatment with increas-

ing concentration of KCl, which disrupts electrostatic interactions, revealed that the salt concentrations that completely removed MHC B resulted in core structures composed of paramyosin and minor associated proteins (Deitiker and Epstein, 1993), with MHC A located in the central region (Epstein et al., 1985). Because hydrophobic interactions are resistant to high salt, this observation is consistent with an important role of increased MHC A coat hydrophobicity in the assembly of this structure. When similar experiments were performed with thick filaments isolated from a mutant that lacks MHC B and thus has MHC A along the entire length of the filament, the central MHC A zone remaining on the core structure was expanded, suggesting that this increased interaction is not restricted to the smaller portion of the filament core where MHC A associates in wild type (Epstein et al., 1986). These experimental observations support our hypothesis that the higher coat hydrophobicity in regions 1 and 2 has functional consequences for MHC A assembly, and suggest that these differences result in more stable association of MHC A with itself and/or with paramyosin.

In addition, genetic experiments indicate that MHC A and paramyosin share a close interaction in vivo, perhaps mediated by their matching hydrophobic coat. Genomic duplications that increase MHC A expression suppress the defects in structure and motility associated with *e73*, a missense allele of paramyosin (Riddle and Brenner, 1978; Waterston et al., 1982; Brown and Riddle, 1985). Homozygous *e73* animals accumulate wild-type levels of paramyosin protein (Waterston et al., 1977), but move poorly and show disrupted muscle structure in which aberrant accumulations of paramyosin appear at the ends of muscle cells (Waterston et al., 1977; Epstein et al., 1987; Gengyo-Ando and Kagawa, 1991). Transgenic copies of MHC A, but not MHC B, can also suppress *e73*, and this A-specific suppression maps to the rod (Hoppe, P.E., and R.H. Waterston, unpublished data). Thus, sequences within the rod allow MHC A to interact strongly enough with paramyosin, such that increased concentrations can drive assembly of the mutant protein into a more wild-type structure.

Given the hypothesis that MHC A and paramyosin share assembly behaviors because of their very similar coat hydrophobicity profiles (Fig. 7 b) and the sequence alignment of paramyosin with the light meromyosin portion of myosin, it is tempting to speculate that myosin and

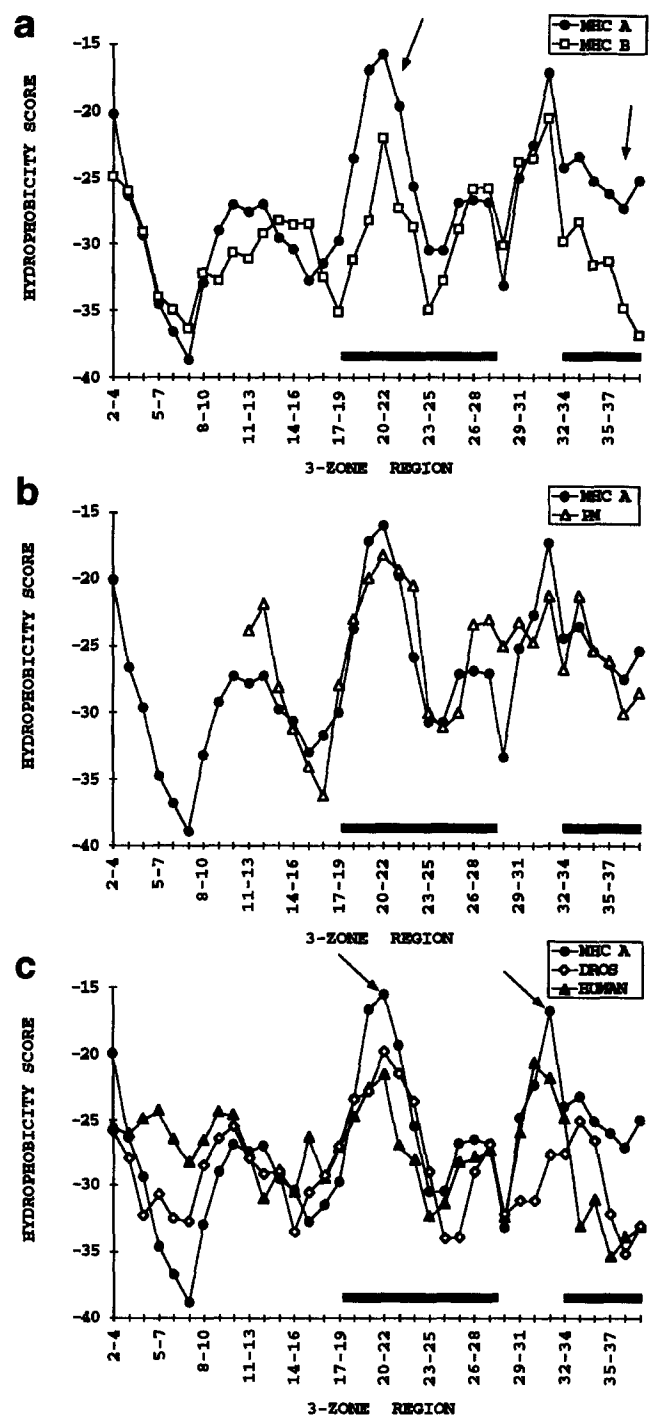


Figure 7. Analysis of coat hydrophobicity along the length of the coiled-coil rod. The hydrophobicity score was calculated by summing the Kyte and Doolittle (1982) scores for all residues in coat positions within each zone, and then determining the average sum within a three-zone interval. ("Zone" refers to the 28-residue repeat defined by McLachlan and Karn [1982]). The three-zone window of sequence was shifted along the length of the rod from zone 2 through 39 in one-zone increments. Similar graphs were obtained using the Engelman et al. (1986) scale. The two horizontal black bars indicate regions of the graph that contain at least one zone of sequence from one of the regions found sufficient for MHC A-specific function. (a) The graphs of the hydrophobicity scores of MHC A and MHC B reveal that coat hydrophobicity varies along the length of the rod, and that the overall pattern of this fluctuation is quite similar in both isoforms. The two areas of the graph where the scores of MHC A and B

most different (arrows) are contained within those regions we have identified as important for thick filament initiation in vivo. (b) Comparison of the hydrophobicity scores of MHC A and paramyosin reveals that the graph of the paramyosin sequence, when aligned by homology, is quite similar to the graph of MHC A, particularly within the regions we have defined as important for A-specific function. (c) Striated muscle myosins from vertebrates and invertebrates show a common overall hydrophobicity profile, particularly in the region from repeat 11 to the end of the rod. The two prominent peaks (arrows) are separated by ~10–11 zones (~280–308 residues or 416–460 Å), which is close to the 430-Å axial spacing that is observed by x-ray diffraction of muscle and associated with the packing of the myosin molecule into the thick filament (Huxley and Brown, 1967).

paramyosin assemble with equivalent staggers. In this light, it is interesting to note that all missense mutations that disrupt paramyosin assembly (Gengyo-Ando and Kagawa, 1991) fall within the regions of paramyosin that are homologous to those we have defined as important for myosin assembly.

What is the Structural Basis for MHC A-specific Function? The *in vitro* and *in vivo* observations discussed above suggest that MHC A forms more stable associations with itself and/or with paramyosin than does MHC B. Previous models of myosin–myosin and myosin–paramyosin interaction in assembly have focused on the matching of charged residues. Recent work examining the solubility of mutant fragments in which the charge pattern had been disrupted indicates, however, that a simple summing of charge–charge interactions along the length of the rod is not sufficient to explain the association of myosin *in vitro* (Atkinson and Stewart, 1992). Below, we present two models in which the relatively greater coat hydrophobicity within MHC A regions 1 and 2 confers A-specific activity by contributing to more stable association of dimers within the thick filament.

In the first model, the increased coat hydrophobicity is critical to filament initiation because it contributes to a stronger interaction between MHC A dimers in antiparallel association; because of reduced coat hydrophobicity, antiparallel MHC B dimers are not sufficiently stable to promote filament assembly. In this light, it is interesting to note that MHC B is by far the least hydrophobic of any myosin in region 2, and is the only myosin known not to be capable of filament initiation. Because the rescue of A-specific function is dominant (that is, a chimeric myosin requires MHC A sequences in either region 1 or region 2, but not both), this model requires that the high hydrophobic content of a single region is sufficient to provide the necessary (although perhaps not optimal) stability. The partial function observed when we divided region 2 sequences (Fig. 3) would then reflect the formation of an antiparallel dimer that was of intermediate stability, close to the threshold level required for filament initiation.

In the second model, A-specific function reflects a tight association of MHC A with paramyosin via a rod surface that supports this interaction. In this case, because the paramyosin hydrophobicity profile is like that of MHC A, the dominance of A-specific rescue can be explained if one requires a good match at only one of the two hydrophobic regions during myosin–paramyosin assembly (e.g., MHC A region 1 with its paramyosin target).

What regions of MHC A or paramyosin contact the residues within regions 1 and 2 in antiparallel alignment? Models based solely upon maximization of charge interactions between antiparallel dimers favor an overlap with zero stagger ($s = 0$) such that the two rods associate along their entire lengths (Kagawa et al., 1989; Fig. 8 *a* in this paper). This stagger results in the association of region 1 of one rod with region 1 of the other, as well as the pairing of the two NH₂-terminal peaks of the hydrophobicity profile (see below). However, it also results in the association of region 2 with sequences in the prehinge region of its partner molecule, the portion of the rod that is most variable in sequence (Dibb et al., 1989) and thought to be only loosely associated with the filament backbone. One in-

triguing alternative is that region 1 of one rod contacts region 2 of its partner molecule during filament initiation. For example, an antiparallel stagger of $s = 452$, which is also favored by charge interaction modeling (Kagawa et al., 1989), results in a direct association between the portions of region 1 and region 2 that are most different in hydrophobicity in MHCs A and B (Fig. 8 *a*). This stagger would produce a tail–tail overlap of 605 residues which, assuming a helical rise of 1.485 Å per residue (Elliot et al., 1968), gives an expected length of 898 Å, or essentially the 900-Å region of overlap measured in type II segments of vertebrate striated muscle myosin in antiparallel association *in vitro* (Harrison et al., 1971).

The $s = 452$ stagger does not involve the NH₂-terminal prehinge portion of the rod, consistent with biochemical studies that demonstrate that the COOH-terminal two thirds of the rod is sufficient to form ordered aggregates, while the prehinge portion is quite soluble and unnecessary (Lowey et al., 1967). Since the prehinge region is in effect the segment lacking in paramyosin, this stagger can be used to explain both myosin and paramyosin antiparallel association (Fig. 8 *a*). While staggers that result in incomplete overlap of the rods might seem at odds with the observed length of the bare zone, which is ~1,600 Å and approximately equal to the length of the myosin rod, subsequent dimer–dimer association in a staggered fashion could produce a filament bare zone of smaller size in a manner similar to that proposed for minifilament assembly in *Acanthamoeba* nonmuscle myosin (Sinard et al., 1989).

Although A-specific function presumably involves antiparallel association, we were interested to determine where regions 1 and 2 fall in the parallel dimer. The $s = 296$ parallel stagger is favored by charge modeling studies, and agrees well with the 430-Å axial spacing detected by x-ray diffraction and EM (McLachlan and Karn, 1982; Kagawa et al., 1989). Pairing of myosin rods with the $s = 296$ stagger results in contact of regions 1 and 2, but the area of overlap does not include the portion of region 1 that differs in hydrophobicity between MHCs A and B (Fig. 8 *c*). Parallel association of paramyosin is postulated to involve the $s = 493$ (720 Å) stagger (Kagawa et al., 1989; Gengyo-Ando and Kagawa, 1991). Interestingly, like the antiparallel $s = 452$ stagger, assembly of myosin dimers with the parallel $s = 493$ stagger results in the association of the two sites that show the greatest difference in coat hydrophobicity in MHCs A and B (Fig. 8 *b*).

The Conserved Hydrophobicity Profile May Be Important for Determining Stagger of Assembly

Given the proposed overall importance of hydrophobicity in stabilizing dimer–dimer associations, we wanted to determine how the characteristic fluctuations in coat hydrophobicity fit with various proposed molecular staggers. As shown in Fig. 7 *c*, the coat hydrophobicity profile, conserved in all striated muscle myosins, contains two prominent peaks. The regions we have defined map near, but do not coincide precisely with these peaks. The more NH₂-terminal peak falls in the portion of region 1 that is most different in hydrophobicity between MHCs A and B. The more COOH-terminal peak falls just outside of region 1.

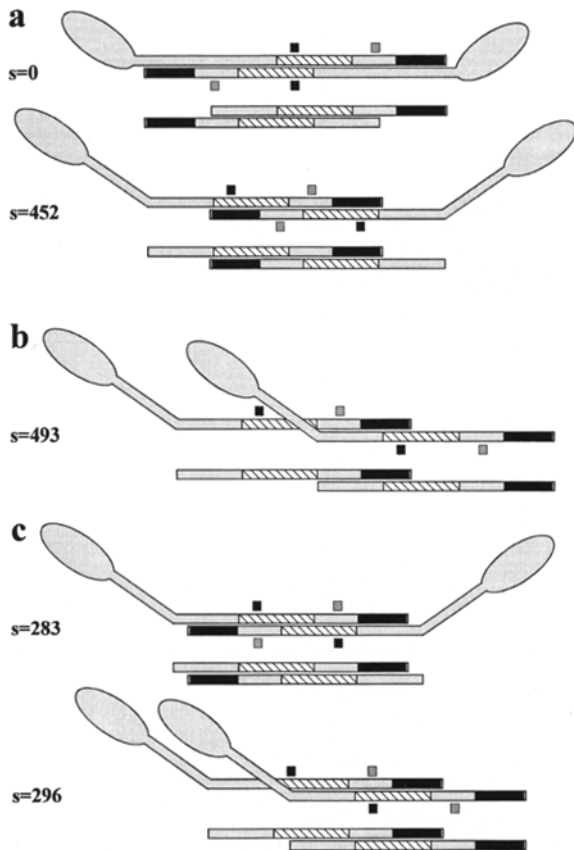


Figure 8. The positions of regions 1 and 2 and the peaks of the coat hydrophobicity profile within dimers associated with different proposed molecular staggers. Charge interaction modeling scores and the computation of staggers are from Kagawa et al. (1989). The hatched and black rod segments represent regions 1 and 2, respectively. The black and grey boxes next to the rod indicate the positions of zones 21 and 31, respectively, which are the zones that fall at the approximate centers of the hydrophobicity peaks (Fig. 7 c) and that have the highest coat hydrophobicity scores in MHC A. (a) Myosin dimers shown associating with two possible antiparallel staggers. The $s = 0$ stagger is favored if charge interactions alone are considered. When paramyosin dimers are drawn with an equivalent stagger, the COOH-terminal region fails to associate with any part of its partner molecule. The $s = 452$ stagger gives relatively high scores in charge modeling of paramyosin-paramyosin and myosin-paramyosin interactions, and results in the direct association of the two rod regions that we have defined as important for filament initiation function. In addition, the stagger works well for the antiparallel association of paramyosin dimers. Of course, other staggers based on these criteria are possible. (b) The $s = 493$ stagger proposed for parallel paramyosin-paramyosin association (Kagawa et al., 1989; Gengyo-Ando and Kagawa, 1991). This parallel stagger results in the association of the rod regions that show the greatest difference in coat hydrophobicity between MHCs A and B, the most NH_2 -terminal part of region 1 with the most COOH-terminal part of region 2. (c) Antiparallel and parallel staggers that pair the two prominent peaks of the hydrophobicity profile, resulting in the direct association of zones 21 and 31. The $s = 283$ antiparallel stagger scores well in charge modeling, and results in a tail-tail overlap of $\sim 1,150 \text{ \AA}$, involving most of light meromyosin, which produces no direct contact between regions 1 and 2. The $s = 296$ parallel stagger (see text) does produce an overlap of regions 1 and 2, but the overlap does not include the portion of region 1 that differs in hydrophobicity between MHCs A and B.

Intriguingly, the $s = 296$ parallel stagger (see above) results in the direct pairing of the more NH_2 -terminal hydrophobicity peak of one rod with the more COOH-terminal peak of its partner (Fig. 8 c). Models of thick filament structure can incorporate parallel staggers of both $s = 296$ (430 \AA) and $s = 493$ (720 \AA) by having, for instance, parallel dimers assemble in one long, single-width subfilament with a 430-\AA stagger, and then having side-to-side association of these subfilaments with the 720-\AA stagger (Chapter 9 in Squire, 1981). The antiparallel stagger $s = 283$ results in a similar pairing of the hydrophobicity peaks of interacting rods (Fig. 8 c), and produces no overlap of the portions of Region 1 and 2 that differ in hydrophobicity in MHCs A and B. Future studies, including the characterization of hydrophobicity profiles of myosins from distantly related nematodes, as well as in vitro mutagenesis experiments in *C. elegans*, may elucidate the role of variations in coat hydrophobicity in both antiparallel and parallel assembly of myosin and paramyosin into the thick filament.

We gratefully acknowledge the scientific input of members of the Waterston and Schedl labs. We owe a special thanks to Bob Barstead for instructing P.E. Hoppe in molecular biological methods. We thank Doug Coulter for scientific discussions, Susan Lowey, Kathy Miller, Chelly Hresko, and Doug Coulter for critical reading of the manuscript, and Tom Blackwell, Haroon Taqi, and Sean Eddy for their work in statistical analysis. We thank Jane Kiff for genomic DNA cassettes. In addition, we are grateful to an anonymous reviewer for corrections to our figures.

This work was supported by U.S. Public Health Service grant GM23833 and a Muscular Dystrophy Association grant awarded to R.H. Waterston, as well as National Research Service Award fellowship 5 F32 GM12412-03 awarded to P.E. Hoppe.

Received for publication 8 May 1996 and in revised form 25 July 1996.

References

- Atkinson, S.J., and M. Stewart. 1992. Molecular interactions in myosin assembly: role of the 28-residue charge repeat in the rod. *J. Mol. Biol.* 226:7-13.
- Brenner, S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics*. 77:71-94.
- Brown, S.J., and D.L. Riddle. 1985. Gene interactions affecting muscle organization in *Caenorhabditis elegans*. *Genetics*. 110:421-440.
- Castellani, L., B.W. Elliott, Jr., and C. Cohen. 1988. Phosphorylatable serine residues are located in a non-helical tailpiece of a catch muscle myosin. *J. Muscle Res. Cell Motil.* 9:533-540.
- Cohen, C., and D.A.D. Parry. 1990. α -helical coiled coils and bundles: how to design an α -helical protein. *Proteins*. 7:1-15.
- Cohen, C., and D.A.D. Parry. 1994. α -helical coiled coils: more facts and better predictions. *Science (Wash. DC)*. 263:488-489.
- Cohen, C., D.E. Lanar, and D.A.D. Parry. 1987. Amino acid sequence and structural repeats in schistosome paramyosin match those of myosin. *Biosci. Rep.* 7:11-16.
- Cotton, R.G.H., N.R. Rodrigues, and R.D. Campbell. 1988. Reactivity of cytosine and thymidine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc. Natl. Acad. Sci. USA*. 85:4397-4401.
- Deitiker, P.R., and H.E. Epstein. 1993. Thick filament substructures in *Caenorhabditis elegans*: evidence for two populations of paramyosin. *J. Cell Biol.* 123:303-311.
- Dibb, N.J., I.N. Maruyama, M. Krause, and J. Karn. 1989. Sequence analysis of the complete *Caenorhabditis elegans* myosin heavy chain gene family. *J. Mol. Biol.* 205:603-613.
- Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. *Am. Stat.* 37:36-48.
- Elliot, A., J. Lowy, D.A.D. Parry, and P.J. Vibert. 1968. Puzzle of the coiled coils in the α -protein paramyosin. *Nature (Lond.)*. 218:656-659.
- Engelman, D.M., T.A. Steitz, and A. Goldman. 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Chem.* 15:321-353.
- Epstein, H.F., R.H. Waterston, and S. Brenner. 1974. A mutant affecting the heavy chain of myosin in *C. elegans*. *J. Mol. Biol.* 90:291-300.
- Epstein, H.F., D.M. Miller, III, I. Ortiz, and G.C. Berliner. 1985. Myosin and paramyosin are organized about a newly identified core structure. *J. Cell Biol.* 100:904-915.

- Epstein, H.F., I. Ortiz, and L.A.T. Mckinnon. 1986. The alteration of myosin compartmentation in specific mutants of *Caenorhabditis elegans*. *J. Cell Biol.* 103:985-993.
- Epstein, H.F., I. Ortiz, and G.C. Berliner. 1987. Assemblages of multiple thick filaments in nematode mutants. *J. Muscle Res. Cell Motil.* 8:527-536.
- Fire, A. 1986. Integrative transformation of *Caenorhabditis elegans*. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2673-2680.
- Fire, A., and R.H. Waterston. 1989. Proper expression of myosin genes in transgenic nematodes. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3419-3428.
- Gengyo-Ando, K., and H. Kagawa. 1991. Single charge change on the helical surface of the paramyosin rod dramatically disrupts thick filament assembly in *Caenorhabditis elegans*. *J. Mol. Biol.* 219:429-441.
- George, E.L., M.B. Ober, and C.P. Emerson, Jr. 1989. Functional domains of the *Drosophila melanogaster* muscle myosin heavy-chain gene are encoded by alternatively spliced exons. *Mol. Cell Biol.* 9:2957-2974.
- Harrison, R.G., S. Lowey, and C. Cohen. 1971. Assembly of myosin. *J. Mol. Biol.* 59:531-535.
- Horton, R.M., H.D. Hunt, S.N. Ho, J.K. Pullen, and L.R. Pease. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene (Amst.)* 77:61-68.
- Huxley, H.E., and W. Brown. 1967. The low-angle x-ray diagram of vertebrate striated muscle and its behavior during contraction and rigor. *J. Mol. Biol.* 30:383-434.
- Kagawa, H., K. Gengyo, A.D. McLachlan, S. Brenner, and J. Karn. 1989. Paramyosin gene (*unc-15*) of *Caenorhabditis elegans*. *J. Mol. Biol.* 207:311-333.
- Kunkel, T.A., J.D. Roberts, and R.A. Zakour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154:367-382.
- Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105-132.
- Lowey, S., L. Goldstein, C. Cohen, and S.M. Luck. 1967. Proteolytic degradation of myosin and the meromyosins by a water-insoluble polyanionic derivative of trypsin. *J. Mol. Biol.* 23:287-304.
- Lu, R.C., and A. Wong. 1985. The amino acid sequence and stability predictions of the hinge region in myosin subfragment 2. *J. Biol. Chem.* 260:3456-3461.
- Mackenzie, J.M., Jr., and H.F. Epstein. 1980. Paramyosin is necessary for determination of nematode thick filament length in vivo. *Cell.* 22:747-755.
- Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
- McLachlan, A.D. 1984. Structural implications of the myosin amino acid sequence. *Annu. Rev. Biophys. Bioeng.* 13:167-189.
- McLachlan, A.D., and J. Karn. 1982. Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle. *Nature (Lond.)* 299:226-231.
- McLachlan, A.D., and M. Stewart. 1975. Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. *J. Mol. Biol.* 98:293-304.
- Mello, C.C., J.M. Kramer, D. Stinchcomb, and V. Ambros. 1991. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:3959-3970.
- Miller, D.M., and I. Maruyama. 1986. The *sup-3* locus is closely linked to a myosin heavy chain gene in *Caenorhabditis elegans*. *UCLA Symp. Mol. Cell Biol. New Ser.* 29:629-638.
- Miller, D.M., I. Ortiz, G.C. Berliner, and H.F. Epstein. 1983. Differential localization of two myosins within nematode thick filaments. *Cell.* 34:477-490.
- Montandon, A.J., P.M. Green, F. Giannelli, and D.R. Bentley. 1989. Direct detection of point mutations by mismatch analysis. Application to haemophilia B. *Nucleic Acids Res.* 17:3347-3358.
- Mooseker, M.S., and R.E. Cheney. 1995. Unconventional myosins. *Annu. Rev. Cell Dev. Biol.* 11:633-675.
- Otsuka, A.J. 1986. *sup-3* suppression affects muscle structure and myosin heavy chain accumulation in *Caenorhabditis elegans*. *UCLA Symp. Mol. Cell Biol. New Ser.* 29:619-628.
- Parry, D.A.D. 1981. Structure of rabbit skeletal myosin: analysis of the amino acid sequences of two fragments from the rod region. *J. Mol. Biol.* 153:459-464.
- Riddle, D.L., and S. Brenner. 1978. Indirect suppression in *Caenorhabditis elegans*. *Genetics.* 89:299-314.
- Schriefer, L.A., and R.H. Waterston. 1989. Phosphorylation of the N-terminal region of *Caenorhabditis elegans* paramyosin. *J. Mol. Biol.* 207:451-454.
- Sinard, J.H., W.F. Stafford, and T.D. Pollard. 1989. The mechanism of assembly of *Acanthamoeba* myosin-II minifilaments: minifilaments assemble by three successive dimerization steps. *J. Cell Biol.* 109:1537-1547.
- Squire, J.M. 1981. The structural basis of muscle contraction. Plenum Publishing Co., New York. 698 pp.
- Waterston, R.H. 1989. The minor myosin heavy chain, mhcA, of *Caenorhabditis elegans* is necessary for the initiation of thick filament assembly. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:3429-3436.
- Waterston, R.H., R.M. Fishpool, and S. Brenner. 1977. Mutants affecting paramyosin in *Caenorhabditis elegans*. *J. Mol. Biol.* 117:679-697.
- Waterston, R.H., D.G. Moerman, D.L. Baillie, and T.R. Lane. 1982. Mutations affecting myosin heavy chain accumulation and function in the nematode *Caenorhabditis elegans*. In *Diseases of the Motor Unit*. D.L. Schotland, editor. Wiley & Sons, New York. 747-760.
- Williams, B.D., B. Schrank, C. Huynh, R. Shownkeen, and R.H. Waterston. 1992. A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. *Genetics.* 131:609-624.
- Wood, W.B. 1988. The Nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 667 pp.